

Characterizing short-term stability for Boolean networks over any distribution of transfer functions

C. Seshadhri,¹ Andrew M. Smith,¹ Yevgeniy Vorobeychik,² Jackson R. Mayo,¹ and Robert C. Armstrong¹

¹*Sandia National Laboratories, P.O. Box 969, Livermore, California 94551-0969, USA*

²*Vanderbilt University, Nashville, TN 37235, USA*

We present a characterization of short-term stability of random Boolean networks under *arbitrary* distributions of transfer functions. Given any distribution of transfer functions for a random Boolean network, we present a formula that decides whether short-term chaos (damage spreading) will happen. We provide a formal proof for this formula, and empirically show that its predictions are accurate. Previous work only works for special cases of balanced families. It has been observed that these characterizations fail for unbalanced families, yet such families are widespread in real biological networks.

INTRODUCTION

Living systems composed of a wide variety of cells, genes, or organs operate with uncanny synchrony and stability, as do numerous engineered and social systems. In a series of seminal papers, Kauffman introduced *Boolean networks* to study such systems: this abstraction involves a network representing connectivity, and a family of Boolean functions determining states of network nodes to model dynamic behavior [1, 2]. Boolean networks have been used to model numerous dynamical systems, including genetic regulatory networks [1] and political systems [3], and have received much theoretical attention [4–12].

A Boolean network has a set of n nodes linked to each other by a directed graph G . Each node i has a Boolean state in $\{-1, +1\}$, an in-degree K_i , and an associated Boolean function $f_i : \{-1, +1\}^{K_i} \rightarrow \{-1, +1\}$, termed *transfer function*. If the state of node i at time t is $x_i(t)$, its state at time $t + 1$ is described by $x_i(t+1) = f_i(x_{i_1}(t), \dots, x_{i_{K_i}}(t))$. For the sake of analysis, it is common to study a randomized ensemble of Boolean networks. The graph G is a directed Erdős-Rényi network, where each vertex i chooses K_i in-neighbors uniformly at random. There is an underlying distribution (or family) of Boolean transfer functions \mathcal{F} . Each vertex i independently chooses the transfer function f_i from \mathcal{F} .

A key parameter of interest is the *short-term stability* of the Boolean network. Specifically, if a single node has its state flipped, does the effect of this perturbation die out (quiescence), exponentially cascade over time (chaos), or is the system right in between (criticality)? There have been numerous empirical and mathematical observations about the characteristics of critical transition points in classes of Boolean networks [4–9, 11–15]. These results require \mathcal{F} to have specific properties: for example, each truth table entry is i.i.d. or that functions are *balanced* (number of $+1$ and -1 outcomes is the same) on average.

These are severe restrictions. Various classes of func-

tions occur naturally in biological and social applications, but do not satisfy either of these conditions. For example, Kauffman proposed a family of *canalyzing functions* [2]. A canalyzing function has at least one input, and one value of that input, that fully determines the output of the function. Kauffman observed that many elements of genetic regulatory systems have like nested canalyzing functions [2, 4, 8, 16]. Previous formal analyses do not yield precise characterizations of short-term stability for such families.

Threshold functions also occur in understanding processes on social and biological networks [17–23]. A threshold function is of the form $f(x_1, x_2, \dots, x_K) = \text{sign}(\sum_i c_i x_i - \Theta)$, where c_i s and Θ are constants. Often there is a bias towards a particular state, so these previous characterizations fail to predict the critical threshold [11].

Our main result gives an exact formula for predicting the short-term dynamics of Boolean networks, for *any* distribution of transfer functions. We stress that our results are for ‘semi-annealed’ setting. Once we choose the topology of the Boolean network and the transfer functions from the appropriate distribution, we assume it is fixed. (We do not change these for each time-step, as in an annealed approximation.) All we need from the topology is a local tree-structure (as proved in [11] and subsequently used in [12]), which is guaranteed with high probability for Erdős-Rényi random graph distributions.

While no previous result provides such a formula, our work is closely related to the following. Mozeika and Saad [10, 14, 15] give a powerful generating function framework for analysis of Boolean networks, but do not characterize short-term stability. Seshadhri et al. [11] introduced the notion of *influence* $I(\mathcal{F})$ of transfer function distribution \mathcal{F} , an easily computable quantity that determines the short-term behavior for a highly restricted class of *balanced* families \mathcal{F} : on average, functions in \mathcal{F} are equally likely to output $+1$ and -1 .

PRELIMINARIES

We are interested in the sensitivity of a Boolean network state $x(t) = \{x_1(t), \dots, x_n(t)\}$ to a small initial perturbation. Formally, consider the following experiment. Suppose that a Boolean network starts from state x , and after t steps reaches a state $F_t(x)$. Now, consider another initial state, $x^{(i)}$ which only differs from x in the i th bit. Let H_t be the expected Hamming distance between $F_t(x)$ and $F_t(x^{(i)})$, where x is drawn from some specified (typically uniform) distribution. How does H_t evolve with time? If H_t can be expressed as $e^{\lambda t}$, then λ is the Lyapunov exponent. If $\lambda < 0$, the boolean network is quiescent; if $\lambda > 0$, the network is chaotic.

We provide some notation and definitions.

- **Biased distributions:** We use \mathcal{D}_ρ to denote the distribution over $\{-1, +1\}$ where the probability of 1 is $(1 + \rho)/2$. We choose this notation because the expected value is exactly ρ , the bias. Abusing notation, for $y \in \{-1, +1\}^K$, we say $y \sim \mathcal{D}_\rho$ when each coordinate of y is chosen i.i.d. from \mathcal{D}_ρ .

- **Imbalance:** The *imbalance* of the Boolean network at time t , denoted by δ_t , is $\sum_{i=1}^n x_i(t)/n$. Informally, this measures the difference between the +1s and -1s in the network. Observe that if the starting state $x(0)$ is chosen from \mathcal{D}_ρ , then $\delta_0 = \rho$.

We use tools from *harmonic analysis of Boolean functions*, pioneered by Kahn, Kalai, and Linial [24]. The convention in this field is that -1 denotes TRUE and $+1$ is FALSE (so multiplication in $\{-1, +1\}$ maps to XOR of $\{0, 1\}$ bits). Consider $f : \{-1, +1\}^K \rightarrow \{-1, +1\}$, where we think of f as one of the transfer functions. The standard representation is as a truth table, with 2^K entries in $\{-1, +1\}$. An alternative representation is as a linear combination of *basis functions*. In the following, we use $y \in \{-1, +1\}^K$ to denote an input to the transfer function. We use $[K]$ for set $\{1, 2, \dots, K\}$, which denotes the input coordinates. Refer to [25] for details on the following.

- **Parity functions:** For any subset S of coordinates in $[K]$, $\prod_{i \in S} y_i$ is the *parity* on S . (For $S = \emptyset$, we set the parity to be 1.)

- **Fourier representation:** Any Boolean function f can be expressed as $f(y) = \sum_{S \subseteq [K]} \hat{f}(S) \prod_{i \in S} y_i$, where $\hat{f}(S)$ are called Fourier coefficients. This expansion represents f as a multilinear polynomial over the Boolean variables y_1, \dots, y_K . It can be shown that $\hat{f}(S) = 2^{-K} \sum_y f(y) \prod_{i \in S} y_i$, the correlation between f and the parity on S . (The Fourier coefficients are the Walsh-Hadamard transform of the truth table.) There are exactly 2^K different Fourier coefficients, one for each subset of the K inputs. For example, consider $K = 2$, and the AND function. A calculation yields $AND(y_1, y_2) = 1/2 + y_1/2 + y_2/2 - y_1 y_2/2$.

- **Level sets of coefficients, σ_r :** Of special interest

is $\sigma_r(f) = \sum_{C: |C|=r} \hat{f}(C)$, where $0 \leq r \leq K$. This is simply the sum of coefficients corresponding to sets of size r . Note that $\sigma_0(f) = \hat{f}(\emptyset) = \sum_y f(y)$. This is exactly the imbalance in the truth table of f .

- **Influence:** For any function f , the influence of the i th variable is denoted $\text{Inf}_i(f) = \Pr_y[f(y) \neq f(y^{(i)})]$ (where the probability is over the uniform distribution and $y^{(i)}$ is obtained by flipping y at the i th bit), and the total influence is $I(f) = \sum_i \text{Inf}_i(f)$. We will define a biased version of this quantity, $\text{Inf}_i(f; \rho) = \Pr_{y \in \mathcal{D}_\rho}[f(y) \neq f(y^{(i)})]$, and analogously $I(f; \rho) = \sum_i \text{Inf}_i(f; \rho)$.

MATHEMATICAL RESULTS

The proofs of our mathematical results are quite involved, and therefore provided in the supplemental material. We can derive closed form expressions for the evolution of δ_t (the expected imbalance at time t) and H_t (the expected Hamming distance at time t after a single bit perturbation).

The evolution of δ_t ($t > 0$) is determined by the level sets of coefficients of the transfer functions. We use $\sigma_r(\mathcal{F}) = \mathbf{E}_{f \sim \mathcal{F}}[\sigma_r(f)]$ and $I(\mathcal{F}; \delta) = \mathbf{E}_{f \sim \mathcal{F}}[I(f; \delta)]$.

Theorem 1 *Let initial state $x(0)$ be chosen from \mathcal{D}_ρ (so $\delta_0 = \rho$). Then δ_t evolves according to the polynomial recurrence $\delta_{t+1} = \sum_{r \geq 0} \sigma_r(\mathcal{F}) \delta_t^r$.*

An equivalent formulation of the recurrence has been derived by the generating function method in Mozeika and Saad [10], though their approach is completely different (they do not show a connection with Fourier coefficients). Our approach proves a clean description of this recurrence, since $\sigma_r(\mathcal{F})$ can be easily computed from \mathcal{F} .

Our main theorem shows how the damage caused by a bit perturbation spreads.

Theorem 2 *Let $\delta_0, \delta_1, \dots$ be as given by Theorem 1. For $t \leq (\log n)/K$, $H_t = \prod_{0 \leq h < t} I(\mathcal{F}; \delta_h)$.*

In many situations, δ_t converges to some δ^* . In that case, $H_t \approx [I(\mathcal{F}; \delta^*)]^t$. The Lyapunov exponent is $\log I(\mathcal{F}; \delta^*)$, so we get a critical point at $I(\mathcal{F}; \delta^*) = 1$. Our formula gives a provable characterization of short-term stability, for *any* transfer function family \mathcal{F} .

Balanced families: As a warmup, we derive previous results that only held for balanced families \mathcal{F} . In such families, the expected difference (over \mathcal{F}) between +1's and -1's in the transfer functions is exactly zero. This contains the classic random families of Kauffman. For such a family, $\sigma_0(\mathcal{F}) = \mathbf{E}_{f \sim \mathcal{F}}[\sigma_0(f)] = 0$. The starting distribution is given by \mathcal{D}_0 , so $\delta_0 = 0$. Regardless of the values of $\sigma_r(\mathcal{F})$ (for $r > 0$), by Theorem 1, $\delta_t = 0$ for all t . Hence, $H_t = [I(\mathcal{F}; 0)]^t$, and $I(\mathcal{F}; 0) = 1$ is the critical threshold. This is exactly the main result of [11].

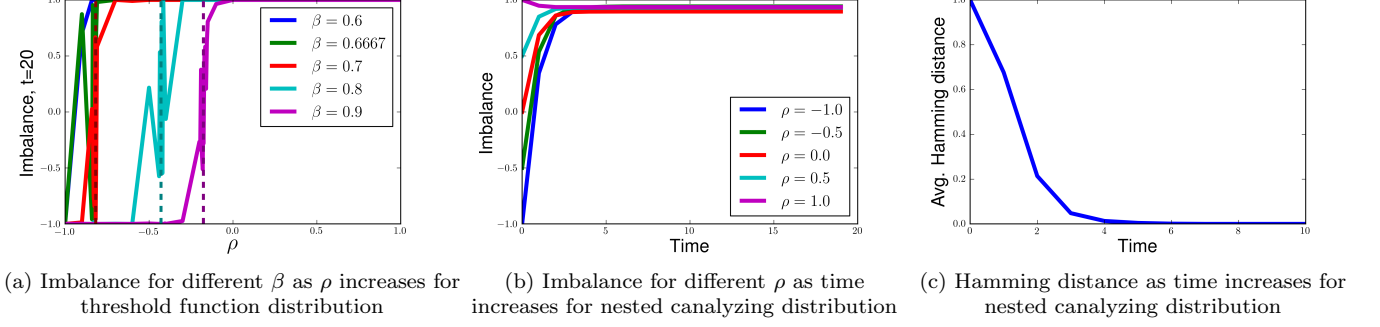


FIG. 1: Experimental results

APPLICATIONS

Mixtures of threshold function families: Threshold functions are commonly used to understand the spread of new ideas/viral propogations in social networks, inspired by pioneering work in sociology [17–19]. Think of two kinds of people (vertices) in a network. Some simply side with the majority of their neighbors. Others are more resistant to change, and only take up a new belief if all their neighbors believe it. We will first demonstrate our theorem on a synthetic distribution inspired by this application. For simplicity of analysis (and to see all the math), set $K = 3$. The majority function is $MAJ(y) = \text{sign}(\sum_i y_i)$ and the AND function $AND(y) = \text{sign}(\sum_i y_i + 2.5)$ (this is -1 iff all inputs are -1). Our distribution \mathcal{F} picks MAJ with probability β and AND with probability $1 - \beta$. How much of the initial network needs to have a new belief for it to propagate through the network? (And how is this sensitive to perturbation?) Formally, what is the dynamics for initial distribution \mathcal{D}_ρ ? Think of a vertex state being -1 (TRUE) if that vertex currently believes the new idea. We start with the Fourier expansions of MAJ and AND .

$$MAJ(y) = \sum_i y_i/2 - y_1 y_2 y_3/2$$

$$AND(y) = 3/4 + \sum_i y_i/4 - \sum_{i \neq j} y_i y_j/4 + y_1 y_2 y_3/4$$

We compute $\sigma_0(\mathcal{F}) = 3(1 - \beta)/4$, $\sigma_1(\mathcal{F}) = 3\beta/2 + 3(1 - \beta)/4 = 3(1 + \beta)/4$, $\sigma_2(\mathcal{F}) = 3(\beta - 1)/4$, and $\sigma_3(\mathcal{F}) = -\beta/2 + (1 - \beta)/4 = (1 - 3\beta)/4$. From Theorem 1,

$$\begin{aligned} \delta_{t+1} = & (1 - 3\beta)\delta_t^3/4 + 3(\beta - 1)\delta_t^2/4 \\ & + 3(1 + \beta)\delta_t/4 + 3(1 - \beta)/4 \end{aligned}$$

Any fixed point is a root of the following polynomial $p(\delta)$ (which basically measures $\delta_{t+1} - \delta_t$). Note that when $p(\delta_t) > 0$, then $\delta_{t+1} > \delta_t$ (and vice versa).

$$\begin{aligned} p(\delta) = & [(1 - 3\beta)\delta^3 + 3(\beta - 1)\delta^2 + (3\beta - 1)\delta + 3(1 - \beta)]/4 \\ = & (\delta - 1)(\delta + 1)[(1 - 3\beta)\delta - 3(1 - \beta)]/4 \end{aligned}$$

This characterizes the limits of δ_t as $t \rightarrow \infty$ (assuming convergence). The first two are trivial roots, since the all -1 s and all $+1$ s states are fixed points imbalances for the Boolean network. The third root $3(1 - \beta)/(1 - 3\beta)$ is a new valid imbalance (in the range $(-1, 1)$) only when $\beta > 2/3$.

Now, we can explain the dynamics. (We ignore the trivial cases $\rho = -1, +1$.)

- $\beta \leq 2/3$: The polynomial $p(z) > 0$ for any $z \in (-1, 1)$. Hence, for any non-trivial starting distribution \mathcal{D}_ρ , the Boolean network converges to the all $+1$ s state. So the new belief will always die out.

- $\beta > 2/3$: There exists a new unstable fixed point for the imbalance at $\delta^* = 3(1 - \beta)/(1 - 3\beta)$. We have $p(z) > 0$ if $z > \delta^*$ and $p(z) < 0$ if $z < \delta^*$. If $\rho > \delta^*$, the eventual state is all $+1$ s. If $\rho < \delta^*$, the eventual state is all -1 s.

To understand the sensitivity to bit flips, it is quite natural that for situations where δ_t converges to -1 or $+1$, the network is insensitive to perturbations. Calculations yield that $\text{Inf}(\mathcal{F}; -1)$ and $\text{Inf}(\mathcal{F}; +1)$ are < 1 . By Theorem 2, the networks are quiescent. At $\rho = \delta^*$, $I(\mathcal{F}; \delta^*) = 3\beta(1 - (\delta^*)^2)/2 + 3(1 - \beta)(1 - \delta^*)^2/4$. By some elementary algebra, $I(\mathcal{F}; \delta^*) > 1$ when $\beta > 2/3$. Hence, for $\rho = \delta^*$, the dynamics are chaotic (again, this is expected).

We performed simulations on Boolean networks with 10^4 nodes. For a given β , we vary the starting distribution ρ and measure the imbalances at $t = 20$. (This was averaged over 1000 runs.) The results are in Figure 1a, where each colored line denotes a different choice of β . The predicted transition of $\delta^* = 3(1 - \beta)/(1 - 3\beta)$ is denoted by the dashed line, coinciding nicely with the experimental transition point. As expected we see some fluctuations (due to chaotic behavior at δ^*) at the transition point.

Nested canalizing functions: For a real application, we consider the nested canalizing functions of [16]. (We provide a full description of this distribution in the supplement.) Previous work suggests that this distribution is reflective of real biological networks and is quiescent. We can use our theorems to validate the quiescence. Let us then consider the polynomial $\delta_{t+1} - \delta_t$. For example at $K = 5$, a technical calculation yields $p(\delta) = -0.001\delta^4 + 0.016\delta^3 - 0.11\delta^2 - 0.69\delta + 0.71$. For $K = 10$, $p(\delta) = -0.007\delta^4 + 0.012\delta^3 - 0.099\delta^2 - 0.7\delta + 0.71$. These polynomials have a single stable root $\delta^* \approx 0.9$ in $[-1, +1]$. Even as K varies, the root is quite stable, so that fixed point imbalance is at least 0.9 *regardless* of the degree distribution.

We perform experiments for varying degree distributions with 10^4 nodes, and varying starting state distributions \mathcal{D}_ρ . (We show only the results for $K = 5$ for space reasons.) In Figure 1b, we plot the imbalance as a function of time for varying ρ . Observe that the imbalance *always* converges to around 0.9. *This means that roughly 90% of the nodes converge to the +1 (FALSE) state.* The influence $I(\mathcal{F}; \delta^*)$ is roughly 0.3, so the network is quiescent. This is validated by the Derrida plot in Figure 1c, which plots average Hamming distance over time (for $\rho = 0$). We observe that the Hamming distance rapidly decays to 0.

ACKNOWLEDGEMENTS

Sandia is a multiprogram laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy under contract DE-AC04-94AL85000.

-
- [1] S. A. Kauffman, "Metabolic stability and epigenesis in randomly connected nets," *Journal of Theoretical Biology*, vol. 22, no. 3, pp. 437–467, 1969.
 - [2] S. A. Kauffman, "The large scale structure and dynamics of gene control circuits: An ensemble approach," *Journal of Theoretical Biology*, vol. 44, no. 1, pp. 167–190, 1974.
 - [3] M. Aldana, S. Coppersmith, and L. P. Kadanoff, "Boolean dynamics with random couplings," in *Perspectives and Problems in Nonlinear Science* (K. E. M. JE, and S. KR, eds.), pp. 23–89, Springer-Verlag, 2003.
 - [4] S. E. Harris, B. K. Sawhill, A. Wuensche, and S. A. Kauffman, "A model of transcriptional regulatory networks based on biases in the observed regulation rules," *Complexity*, vol. 7, no. 4, pp. 23–40, 2002.
 - [5] I. Shmulevich and S. A. Kauffman, "Activities and sensitivities in boolean network models," *Physical Review Letters*, vol. 93, no. 4, p. 048701, 2004.
 - [6] A. A. Moreira and L. A. N. Amaral, "Canalizing kauffman networks: Nonergodicity and its effect on their critical behavior," *Physical Review Letters*, vol. 94, p. 218702, 2005.
 - [7] M. Aldana and P. Cluzel, "A natural class of robust networks," *Proceedings of the National Academy of Sciences*, vol. 100, no. 15, pp. 8710–8714, 2003.
 - [8] S. A. Kauffman, C. Peterson, B. Samuelsson, and C. Troein, "Random boolean network models and the yeast transcriptional network," *Proceedings of the National Academy of Sciences*, vol. 100, no. 25, pp. 14796–14799, 2003.
 - [9] I. Shmulevich, H. Lahdesmaki, E. R. Dougherty, J. Astola, and W. Zhang, "The role of certain Post classes in boolean network models of genetic networks," *Proceedings of the National Academy of Sciences*, vol. 100, no. 19, pp. 10734–10739, 2003.
 - [10] A. Mozeika and D. Saad, "Dynamics of Boolean networks: An exact solution," *Physical Review Letters*, vol. 106, p. 214101, 2011.
 - [11] C. Seshadhri, Y. Vorobeychik, J. R. Mayo, R. C. Armstrong, and J. R. Ruthruff, "Influence and dynamic behavior in random Boolean networks," *Physical Review Letters*, vol. 107, p. 108702, 2011.
 - [12] S. Squires, E. Ott, and M. Girvan, "Dynamic instability in Boolean networks as a percolation problem," *Physical Review Letters*, vol. 109, p. 085701, 2012.
 - [13] J. Kesseli, P. Ramo, and O. Yli-Harja, "On spectral techniques in analysis of boolean networks," *Physica D: Nonlinear Phenomena*, vol. 206, no. 1-2, pp. 49–61, 2005.
 - [14] A. Mozeika, D. Saad, and J. Raymond, "Computing with noise - phase transitions in boolean formulas," *Phys. Rev. Lett.*, vol. 103, p. 248701, 2009.
 - [15] A. Mozeika, D. Saad, and J. Raymond, "Noisy random boolean formulae - a statistical physics perspective," *Phys. Rev. E*, vol. 82, p. 041112, 2010.
 - [16] S. A. Kauffman, C. Peterson, B. Samuelsson, and C. Troein, "Genetic networks with canalizing boolean rules are always stable," *Proceedings of the National Academy of Sciences*, vol. 101, no. 49, pp. 17102–17107, 2004.
 - [17] T. C. Schelling, *Micromotives and Macrobehavior*. Norton, 1978.
 - [18] M. Granovetter, "Threshold models of collective behavior," *American Journal of Sociology*, vol. 83, no. 6, pp. 1420–1443, 1978.
 - [19] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *SIGKDD conference on Knowledge discovery and data mining*, pp. 137–146, 2003.
 - [20] F. Li, T. Long, Y. Lu, Q. Ouyang, and C. Tang, "The yeast cell-cycle network is robustly designed," *Proceedings of the National Academy of Sciences*, vol. 101, pp. 4781–4786, 2004.
 - [21] M. Davidich and D. Bornholdt, "Boolean network model predicts cell-cycle sequence of fission yeast," *PLoS One*, vol. 3, p. e1672, 2008.
 - [22] J. G. T. Zaudo, M. Aldana, and G. Martinez-Mekler, "Boolean threshold networks: Virtues and limitations for biological modeling," *Information Processing and Biological Systems*, vol. 11, pp. 113–151, 2011.
 - [23] V. Tran, M. N. McCall, H. R. McMurray, and A. Al-mudevar, "On the underlying assumptions of threshold boolean networks as a model for genetic regulatory network behavior," *Frontiers in Genetics*, vol. 4, p. 263, 2013.
 - [24] J. Kahn, G. Kalai, and N. Linial, "The influence of variables on boolean functions," in *Twenty-Ninth Symposium on the Foundations of Computer Science*, pp. 68–80, 1988.

[25] R. O'Donnell, *Analysis of Boolean Functions*. Cambridge University Press, 2014.

Supplemental Material

Preliminaries and notation

For convenience, we state and formalize many of the basic concepts already introduced in the main body.

For $\rho \in [-1, 1]$, we define a biased distribution \mathcal{D}_ρ on \mathcal{B} as follows. The probability of $+1$ is $(1 + \rho)/2$ and that of -1 is $(1 - \rho)/2$. Note that expectation is exactly ρ . We sometimes abuse notation and use \mathcal{D}_ρ to denote the product distribution over n bits. The uniform distribution is given by \mathcal{D}_0 .

We assume that there is a distribution \mathcal{T} on transfer functions. Formally, this is a union of distributions \mathcal{T}_d , where this family only contains boolean functions that take d inputs. For each vertex v with indegree d , we first choose an independent function $f_v(y_1, y_2, \dots, y_d)$ from \mathcal{T}_d . Randomly permute the in-neighbors of i to get a list v_1, v_2, \dots, v_d . Assign the vertex v_j to input y_j of ϕ_i . This gives us the transfer function for vertex i .

A convenient method for ignoring varying degrees is the following. We assume that each vertex has an indegree of K , with neighbors chosen randomly as before. Any function ϕ with less than $d < K$ inputs can be extended to have K inputs, where ϕ does not depend on the new $K - d$ inputs. We now apply the same construction, where there is a single distribution \mathcal{T} over input functions.

For a boolean network \mathcal{N} , we use $f_t(x)$ to denote the total state after t steps starting with an initial state x . We use $f_{v,t}(x)$ to denote the (boolean) state at the vertex v . Our aim is to understand $H_t = (1/n) \sum_{i=1}^n \mathbf{E}_{x \in \mathcal{D}_\rho} [f_t(x) - f_t(x^{(i)})]$. Meaning, we look at the expected Hamming distance over the starting state x for a random bit flip. As proven in previous work, this is the same as $\frac{1}{n} \sum_{1 \leq u, v \leq n} \text{Inf}_u(f_{v,t}; \rho)$. This is the average value (over all vertices v) of $\sum_u \text{Inf}_u(f_{v,t}; \rho)$. Since the construction of boolean networks is random where all vertices are symmetric, in expectation, all these influence sums are the same. Hence, we will fix a single vertex and focus on this sum.

Fourier Analysis of Boolean Functions

We will focus on functions of the form $f : \{-1, +1\}^K \rightarrow \{-1, +1\}$. We think of a function as a vector in \mathbb{R}^{2^K} , which is just an explicit representation of the truth table. The Fourier basis for Boolean functions (also called the Walsh-Hadamard basis) provides an alternate basis to represent functions.

Definition 3 • Let $S \subseteq [K]$. The parity on S is the function $\chi_S(y) = \prod_{i \in S} y_i$. Conventionally, the

function χ_\emptyset is a constant function that takes value +1.

- For $S \subseteq [K]$, define $\hat{f}(S) := \mathbf{E}_{x \in \mathcal{D}_0}[f(y)\chi_S(y)] = 2^{-K} \sum_x f(y)\chi_S(y)$.

The fundamental theorem is that the parities form an orthonormal basis for functions f on the \mathcal{B}^d . This gives the *Fourier expansion* of f .

Theorem 4 Every function $f : \{-1, +1\}^d \rightarrow \mathbb{R}$ is uniquely expressible as a linear combination of the parity functions. Formally, $f = \sum_{S \subseteq [K]} \hat{f}(S)\chi_S$.

The influences are fundamentally connected to the Fourier expansion.

Proposition 5 The value of $\text{Inf}_i(f; \rho)$ is equal to the following three expressions.

- $(1/4)\mathbf{E}_{x \sim \mathcal{D}_\rho}[(f(y) - f(y^{(i)}))^2]$
- $\mathbf{E}_{x \sim \mathcal{D}_\rho}[(\sum_{S \ni i} \hat{f}(S)\chi_{S \setminus i}(y))^2]$

Proof: Since the probability distribution is always \mathcal{D}_ρ , we drop the subscript $x \sim \mathcal{D}_\rho$. We have $\text{Inf}_i(f; \rho) = \Pr[f(y) \neq f(y^{(i)})]$. Observe that $(f(y) - f(y^{(i)}))^2 = 4$ if $f(y) \neq f(y^{(i)})$ and zero otherwise. Hence, $4 \cdot \text{Inf}_i(f; \rho) = \mathbf{E}[(f(y) - f(y^{(i)}))^2]$. We expand this expression.

$$\begin{aligned} 4 \cdot \text{Inf}_i(f; \rho) &= \mathbf{E}[(f(y) - f(y^{(i)}))^2] \\ &= \mathbf{E}\left[\left(\sum_S \hat{f}(S)\left(\prod_{j \in S} y_j - \prod_{j \in S} y_j^{(i)}\right)\right)^2\right] \\ &= \mathbf{E}\left[\left(\sum_{S \ni i} \hat{f}(S)(y_i - y_i^{(i)}) \prod_{j \in S \setminus i} y_j\right)^2\right] \quad (\text{since for } j \neq i, y_j = y_j^{(i)}) \\ &= 4\mathbf{E}\left[\left(\sum_{S \ni i} \hat{f}(S) \prod_{j \in S \setminus i} y_j\right)^2\right] \quad (\text{since } |y_i - y_i^{(i)}| = 2) \end{aligned}$$

□

Deriving the recurrences

Fix a vertex v . Let us consider the function $f_{v,t}$ for small $t \ll \log n$. Previous work tells us that we can assume (asymptotically) this is a rooted tree [11]. We use $N_{\leq t, v}^-$ to denote the t -step in-neighborhood of vertex i .

Claim 6 Fix a vertex v and let $t \leq (\log n)/(4K)$. The probability that the subgraph induced by $N_{\leq t, v}^-$ is a directed tree is at least $1 - 1/\sqrt{n}$.

The distribution \mathcal{B}_t : We define a distribution on Boolean networks that runs for t steps on rooted trees with height t . This captures the t -neighborhood of v

based on Claim 6. We take a K -ary directed tree rooted at v of depth t , with edges pointing towards the root v . For every internal node u , we choose a transfer function ϕ_u distributed according to \mathcal{F} . The leaves of the tree are the input nodes, collectively denoted as x . We will set the state at leaf nodes from the distribution \mathcal{D}_ρ . So $\delta_0 = \rho$ is the initial imbalance.

The Boolean network runs for t steps to yield the state at the root. Observe that for a vertex u at height h , only the function $f_{u,h}$ is defined.

We will use v_1, v_2, \dots to denote the children of v . The Fourier expansion yields the following claim. This innocuous statement is the heart of the analysis.

Claim 7 $f_{v,t} = \sum_{A \subseteq [K]} \widehat{\phi}_v(A) \prod_{i \in A} f_{v_i, t-1}$

Proof: Suppose the state at v_i is y_i . The state at v is determined by applying the transfer function ϕ_v on the states (y_1, y_2, \dots, y_K) . Using the Fourier expansion of ϕ_v , we get the state at v is $\sum_{A \subseteq [K]} \widehat{\phi}_v(A) \prod_{i \in A} y_i$. The state y_i is given by the function $f_{v_i, t-1}$, and the state at v is $f_{v,t}$. □

The imbalance recurrence

We derive a polynomial recurrence for δ_t , the expected imbalance at a vertex after t steps. We have $\delta_t = \mathbf{E}_{\mathcal{B}_t}[\mathbf{E}_{x \in \mathcal{D}_\rho}[f_{v,t}(x)]]$. For any r , remember that $\sigma_r = \mathbf{E}_{\phi \sim \mathcal{F}}[\sum_{C: |C|=r} \widehat{\phi}(C)]$.

Theorem 8 Let δ_t be the expected imbalance at time t . For $t \geq 1$, δ_t evolves according to the following iterated polynomial map.

$$\delta_t = \sum_{r \geq 0} \sigma_r \delta_{t-1}^r \quad (1)$$

Proof: We take expectations of the formula in Claim 7. (Verbal explanation follows.)

$$\begin{aligned} \delta_t &= \mathbf{E}_{x, \mathcal{B}_t}[f_{v,t}(y)] = \mathbf{E}_{x, \mathcal{B}_t}\left[\sum_{A \subseteq [K]} \widehat{\phi}_v(A) \prod_{i \in A} f_{v_i, t-1}(y)\right] \\ &= \sum_{A \subseteq [K]} \mathbf{E}_{x, \mathcal{B}_t}\left[\widehat{\phi}_v(A) \prod_{i \in A} f_{v_i, t-1}(y)\right] \\ &= \sum_{A \subseteq [K]} \mathbf{E}_{\mathcal{F}}[\widehat{\phi}(A)] \prod_{i \in A} \mathbf{E}_{x, \mathcal{B}_{t-1}}[f_{v_i, t-1}(y)] \end{aligned}$$

The second line is just linearity of expectation. The final line is obtained through independence. Note that ϕ_v is independent of the Boolean networks rooted at the v_i s. These Boolean networks are also independent of each other. Hence, the expectation of the product is the product of expectations. The function ϕ_v is a random function ϕ chosen from \mathcal{F} . Because of the recursive construction, the distribution of \mathcal{B}_t rooted at v induces the

distribution of \mathcal{B}_{t-1} rooted at the v_i s. Now, observe that $\mathbf{E}_{x, \mathcal{B}_{t-1}}[f_{v_i, t-1}(y)] = \delta_{t-1}$.

Plugging this in and collecting all terms corresponding to sets of the same size,

$$\begin{aligned}\delta_t &= \sum_{A \subseteq [K]} \delta_{t-1}^{|A|} \mathbf{E}_{\mathcal{F}}[\hat{\phi}(A)] \\ &= \sum_{r \geq 0} \delta_{t-1}^r \sum_{A: |A|=r} \mathbf{E}_{\mathcal{F}}[\hat{\phi}(A)] = \sum_{r \geq 0} \sigma_r \delta_{t-1}^r\end{aligned}$$

□

The spreading of perturbations

We focus on $I_t(\rho_0)$, the expected average (over all nodes) influence of a node at t -steps, when the initial distribution is \mathcal{D}_{ρ_0} . By the tree approximation, it suffices to focus on the node v and consider the distribution \mathcal{B}_t . We can express $I_t(\rho_0)$ as follows.

By the tree approximation, $H_t = \mathbf{E}_{\mathcal{B}_t}[\sum_{\ell} \text{Inf}_{\ell}(f_{v,t}; \rho)]$ (where ℓ is over all leaves). In words, we look at the ρ -biased influence summed over all leaves. For convenience, we will drop the time/height subscript and simply write f_u instead of $f_{u,h}$.

Theorem 9 $H_t = \prod_{0 \leq h < t} I(\mathcal{F}; \delta_h)$

Proof: Partition the leaves into subsets S_1, S_2, \dots, S_K , where S_i contains all leaves that are descendants of v_i . Focus on a leaf $\ell \in S_1$. By [Prop. 5](#) and [Claim 7](#),

$$\begin{aligned}\mathbf{E}_{\mathcal{B}_t}[\text{Inf}_{\ell}(f_v; \rho)] &= (1/4) \mathbf{E}_{\mathcal{B}_t, x \sim \mathcal{D}_{\rho}}[(f_v(x) - f_v(x^{(\ell)}))^2] \\ &= (1/4) \mathbf{E}_{\mathcal{B}_t, x \sim \mathcal{D}_{\rho}}\left[\left\{\sum_A \hat{\phi}_v(A) \left(\prod_{i \in A} f_{v_i}(x) - \prod_{i \in A} f_{v_i}(x^{(\ell)})\right)\right\}^2\right]\end{aligned}$$

Observe that for $i \neq 1$, $f_{v_i}(x) = f_{v_i}(x^{(\ell)})$. (This is because ℓ is not in the subtree of v_i .) In the summation above, only the terms corresponding to $A \ni 1$ are non-zero. Expanding further,

$$\begin{aligned}&\left\{\sum_{A \ni 1} \hat{\phi}_v(A) \left(\prod_{i \in A} f_{v_i}(x) - \prod_{i \in A} f_{v_i}(x^{(\ell)})\right)\right\}^2 \\ &= \left\{\sum_{A \ni 1} \hat{\phi}_v(A) \left(\prod_{\substack{i \in A \\ i \neq 1}} f_{v_i}(x)\right) (f_{v_1}(x) - f_{v_1}(x^{(\ell)}))\right\}^2 \\ &= (f_{v_1}(x) - f_{v_1}(x^{(\ell)}))^2 \left\{\sum_{A \ni 1} \hat{\phi}_v(A) \prod_{\substack{i \in A \\ i \neq 1}} f_{v_i}(x)\right\}^2\end{aligned}$$

Each f_{v_i} is defined over disjoint parts of the underlying tree with disjoint inputs. Hence, when we take the expectation $\mathbf{E}_{\mathcal{B}_t, x}$ over the product, we get the product of expectations. Moreover, $(1/4) \mathbf{E}_{\mathcal{B}_t, x \sim \mathcal{D}_{\rho}}[(f_{v_1}(x) - f_{v_1}(x^{(\ell)}))^2]$ is exactly $\mathbf{E}_{\mathcal{B}_{t-1}}[\text{Inf}_{\ell}(f_{v_1}; \rho)]$.

The random variable $f_{v_i}(x)$ is in $\{-1, +1\}$ and $\mathbf{E}_{\mathcal{B}_{t-1}, x \sim \mathcal{D}_{\rho}}[f_{v_i}(x)] = \delta_{t-1}$. Hence, it is distributed as $\mathcal{D}_{\delta_{t-1}}$. Taking expectations over \mathcal{B}_{t-1}, x , setting $y_i = f_{v_i}(x)$ and [Prop. 5](#),

$$\begin{aligned}\mathbf{E}_{\mathcal{B}_t, x \sim \mathcal{D}_{\rho}}\left[\left\{\sum_{A \ni 1} \hat{\phi}_v(A) \prod_{\substack{i \in A \\ i \neq 1}} f_{v_i}(x)\right\}^2\right] \\ &= \mathbf{E}_{\phi \sim \mathcal{F}, y \sim \mathcal{D}_{\delta_{t-1}}}\left[\left\{\sum_{A \ni 1} \hat{\phi}_v(A) \prod_{i \in A \setminus 1} y_i\right\}^2\right] \\ &= \mathbf{E}_{\phi \sim \mathcal{F}}[\text{Inf}_1(\phi; \delta_{t-1})]\end{aligned}$$

In general, for $\ell \in S_i$, we get $\mathbf{E}_{\mathcal{B}_t}[\text{Inf}_{\ell}(f_v; \rho)] = \mathbf{E}_{\mathcal{F}}[\text{Inf}_i(\phi; \delta_{t-1})] \mathbf{E}_{\mathcal{B}_{t-1}}[\text{Inf}_{\ell}(f_{v_i}; \rho)]$. We combine all our observations.

$$\begin{aligned}H_t &= \sum_{\ell} \mathbf{E}_{\mathcal{B}_t}[\text{Inf}_{\ell}(f_v; \rho)] \\ &= \sum_{i=1}^K \sum_{\ell \in S_i} \mathbf{E}_{\mathcal{B}_t}[\text{Inf}_{\ell}(f_v; \rho)] \\ &= \sum_{i=1}^K \mathbf{E}_{\mathcal{F}}[\text{Inf}_i(\phi; \delta_{t-1})] \sum_{\ell \in S_i} \mathbf{E}_{\mathcal{B}_{t-1}}[\text{Inf}_{\ell}(f_{v_i}; \rho)] \\ &= \sum_{i=1}^K \mathbf{E}_{\mathcal{F}}[\text{Inf}_i(\phi; \delta_{t-1})] \mathbf{E}_{\mathcal{B}_{t-1}}\left[\sum_{\ell \in S_i} \text{Inf}_{\ell}(f_{v_i}; \rho)\right] \\ &= H_{t-1} \sum_{i=1}^K \mathbf{E}_{\mathcal{F}}[\text{Inf}_i(\phi; \delta_{t-1})] \\ &= H_{t-1} \cdot I(\mathcal{F}; \delta_{t-1})\end{aligned}$$

Uncoiling the recurrence yields the theorem. □

Nested canalizing functions

For completeness, we describe this distribution. Fix positive integer α and a series of canalizing input values c_1, c_2, \dots, c_K and $d_1, d_2, \dots, d_K, d_{def}$ (where each of these is in $\{-1, +1\}$). The function is defined as follows:

$$f(x) = \begin{cases} d_1 & \text{if } y_1 = d_1 \\ d_2 & \text{if } y_1 \neq d_1 \text{ and } y_2 = d_2 \\ \vdots & \\ d_K & \text{if } y_1 \neq d_1, y_2 \neq d_2, \dots, y_{K-1} \neq d_{K-1} \text{ and } y_K = d_K \\ d_{def} & \text{otherwise} \end{cases}$$

For any parameter $\alpha > 0$, the distribution is given by $\Pr[c_i = -1] = \Pr[d_i = -1] = \exp(-\alpha/2^i)/(1 + \exp(-\alpha/2^i))$. Kauffman et al suggest that $\alpha = 7$ is reflective of real biological networks, and corresponding boolean networks are quiescent.